Department of EECS - University of California at Berkeley

EECS 126 - Probability and Random Processes - Spring 2007

Midterm 2: 4/03/2007

SOLUTIONS

---

1. (20%)

The random variables X, Y are independent. X is uniformly distributed in $[0, 1]$ and Y is exponentially distributed with mean 1, so that $P(Y > y) = e^{-y}$ for $y \geq 0$. Let $Z = \min\{X, Y\}$. Calculate the mean and the variance of Z.

Hint: You may need the following intermediate results. Let $a_n = \int_0^1 x^n e^{-x} dx$ for $n \geq 0$. One finds that $a_0 = 1 - e^{-1}, a_1 = 1 - 2e^{-1}, a_2 = 2 - 5e^{-1}, a_3 = 6 - 16e^{-1}$.

We have $P(Z > x) = P(X > x, Y > x) = P(X > x)P(Y > x) = (1 - x)e^{-x}$ for $x \in [0, 1]$. The density f_Z of Z is thus given by

$$f_Z(x) = -\frac{d}{dx} P(Z > x) = e^{-x} + (1 - x)e^{-x} = 2e^{-x} - xe^{-x}, \text{ for } 0 \leq x \leq 1.$$

Now,

$$E(Z) = \int_0^1 x f_Z(x) dx = 2\int_0^1 xe^{-x} dx - \int_0^1 x^2 e^{-x} dx = 2a_1 - a_2 = 2 - 4e^{-1} - [2 - 5e^{-1}] = e^{-1} \approx 0.37$$

and

$$\begin{aligned} E(Z^2) &= \int_0^1 x^2 f_Z(x) dx = 2\int_0^1 x^2 e^{-x} dx - \int_0^1 x^3 e^{-x} dx = 2a_2 - a_3 = 2[2 - 5e^{-1}] - [6 - 16e^{-1}] \\ &= -2 + 6e^{-1}, \end{aligned}$$

so that

$$\text{var}(Z) = -2 + 6e^{-1} - e^{-2} \approx 0.07.$$

2. (15%)

The random variables X and Y are as in Problem **1**.

a) Calculate $f_V(v)$ where $V = X + Y$.

b) Calculate $E(V)$ and $\text{var}(V)$.

---

a) To find the density of V we note that, if $v < 1$, then

$$P(V > v) = P(X > v) + \int_0^v f_X(x)P(Y > v-x)dx = 1-v+\int_0^v e^{x-v}dx = 1-v+(1-e^{-v}) = 2-v-e^{-v}.$$

Also, if $v > 1$,

$$P(V > v) = \int_0^1 f_X(x)P(Y > v - x)dx = \int_0^1 e^{x-v}dx = e^{-v}[e - 1].$$

We find the density of V by differentiating and we get

$$f_V(v) = \begin{cases} 1 - e^{-v}, & \text{for } v \leq 1 \\ e^{-v}(e - 1), & \text{for } v > 1. \end{cases}$$

b) We find

$$E(V) = E(X) + E(Y) = \frac{1}{2} + 1 = 1.5$$

and

$$\text{var}(V) = \text{var}(X) + \text{var}(Y) = E(X^2) - E(X)^2 + E(Y^2) - E(Y^2) = \frac{1}{3} - \frac{1}{4} + 2 - 1 = \frac{13}{12}$$

since

$$E(Y^2) = \int_0^\infty y^2 e^{-y}dy = -\int_0^\infty y^2 de^{-y} = \int_0^\infty 2ye^{-y}dy = 2.$$

3. (15%)

Let X_1, X_2, X_3 be independent $N(0,1)$ random variables. Calculate $E[X_1|2X_1 + X_2, X_2 + 3X_3]$.

With $X = X_1, Y_1 = 2X_1 + X_2, Y_2 = X_2 + 3X_3$, and $\mathbf{Y} = (Y_1, Y_2)^T$, we find

$$\Sigma_{X\mathbf{Y}} = E[X(Y_1, Y_2)] = [2, 0],$$

and

$$\Sigma_{\mathbf{Y}} = E[\mathbf{Y}\mathbf{Y}^T] = \begin{bmatrix} 5 & 1 \\ 1 & 10 \end{bmatrix}.$$

Hence,

$$\begin{aligned} &E[X_1|2X_1 + X_2, X_2 + 3X_3] = E[X|\mathbf{Y}] = \Sigma_{X\mathbf{Y}}\Sigma_{\mathbf{Y}}^{-1}\mathbf{Y} \\ &= [2, 0]\begin{bmatrix} 5 & 1 \\ 1 & 10 \end{bmatrix}^{-1}\mathbf{Y} = [2, 0]\frac{1}{49}\begin{bmatrix} 10 & -1 \\ -1 & 5 \end{bmatrix}\mathbf{Y} = \frac{1}{49}[20, -2]\mathbf{Y} = \frac{1}{49}[20Y_1 - 2Y_2]. \end{aligned}$$

4. (20%)

The random variables $X, Z_1, Z_2, \ldots$ are independent; X is $N(0, \sigma^2)$ and $Z_n = N(0, u^2)$ for $n \geq 1$. Let $\hat{X}_n = E[X|Y_1, \ldots, Y_n]$ where $Y_k = X + Z_k, k = 1, \ldots, n$. You must choose the number n of measurements so that $P(|X - \hat{X}_n| > 0.1) < 5\%$.

a) Find $\hat{X}_n$.

b) Use Chebyshev's inequality to estimate the smallest value of n you should use if $\sigma^2 = 4, u^2 = 1$ so that $P(|X - \hat{X}_n| > 0.1) < 5\%$.

c) Use the Gaussian distribution to estimate the smallest value of n you should use if $\sigma^2 = 4, u^2 = 1$ so that $P(|X - \hat{X}_n| > 0.1) < 5\%$.

Note: Here are some potentially useful values: $P(|N(0,1)| > 1.64) = 10\%, P(|N(0,1)| > 1.96) = 5\%, P(|N(0,1)| > 2.58) = 1\%$.

---

a) We know, by symmetry, that

$$\hat{X}_n = a(Y_1 + \cdots + Y_n).$$

Also, a should be such that $X - \hat{X}_n$ is orthogonal to each Y_k. Writing that $X - \hat{X}_n \perp Y_1$ we find

$$0 = E((X - a(nX + Z_1 + \cdots + Z_n))(X + Z_1)) = (1 - an)\sigma^2 - au^2,$$

so that

$$a = \frac{\sigma^2}{u^2 + n\sigma^2}.$$

b) We find that

$$E((X - \hat{X})^2) = \text{var}((1 - an)X - aZ_1 - \cdots - aZ_n) = (1 - an)^2\sigma^2 + na^2u^2$$
$$= \frac{u^4\sigma^2}{(u^2 + n\sigma^2)^2} + \frac{nu^2\sigma^4}{(u^2 + n\sigma^2)^2} = \frac{u^2\sigma^2}{u^2 + n\sigma^2}.$$

Using Chebyshev's inequality, we have

$$P(|X - \hat{X}_n| > \epsilon) \leq \frac{E((X - \hat{X})^2)}{\epsilon^2} = \frac{u^2\sigma^2}{\epsilon^2(u^2 + n\sigma^2)}.$$

With $\sigma^2 = 4, u^2 = 1, \epsilon = 0.1$, we see that we need

$$\frac{4}{0.01(1 + 4n)} \leq 0.05,$$

which implies $n \geq 2,000$.

c) We know that $X - \hat{X}^2 = N(0, b_n^2)$ where $b_n^2 = \frac{u^2\sigma^2}{u^2 + n\sigma^2}$. Now,

$$P(|N(0, b_n^2)| > \epsilon) = P(|N(0,1)| > \frac{\epsilon}{b_n}) \leq 0.05 \text{ if } \frac{\epsilon}{b_n} \geq 1.96.$$

Hence, we need $\epsilon / b_n \geq 1.96$, or $b_n \leq \epsilon/1.96$, or $b_n^2 \leq \epsilon^2/(1.96)^2 \approx \epsilon^2/4$. That is,

$$\frac{4}{1 + 4n} \leq \frac{0.01}{4},$$

so that $n \geq 400$. The example shows that the Chebyshev estimate is a bit pessimistic.

5. (15%)

Assume that Z_1, Z_2 are independent, zero mean, and with respective variances 1, 4. The random variable X is independent of $\{Z_1, Z_2\}$ and is equally likely to be equal to -1 or $+1$.

a) Find $\hat{X} = MAP[X|X + Z_1, X + Z_2]$.

b) Find $P(\hat{X} \neq X)$.

Hint: Here are some values of $F_W(x)$ for $W = N(0, 1)$ that you may need:

x	-1.6	-1.5	-1.4	-1.3	-1.2	-1.1	-1.0	-0.9	-0.8	-0.7	-0.6
$F_W(x)$	0.055	0.067	0.081	0.097	0.115	0.136	0.159	0.184	0.212	0.242	0.274

---

a) Let $Y_1 = X + Z_1$ and $Y_2 = X + Z_2$. We find

$$f_{[\mathbf{Y}|X]}[\mathbf{y}|1] = \frac{1}{\sqrt{2\pi}} e^{-(y_1-1)^2/2} \frac{1}{\sqrt{2\pi 4}} e^{-(y_2-1)^2/8}$$

and

$$f_{[\mathbf{Y}|X]}[\mathbf{y}|-1] = \frac{1}{\sqrt{2\pi}} e^{-(y_1+1)^2/2} \frac{1}{\sqrt{2\pi 4}} e^{-(y_2+1)^2/8}.$$

Hence,

$$L(\mathbf{y}) = \frac{f_{[\mathbf{Y}|X]}[\mathbf{y}|1]}{f_{[\mathbf{Y}|X]}[\mathbf{y}|-1]} = \exp\{2y_1 + \frac{1}{2}y_2\}.$$

Accordingly,

$$MAP[X|\mathbf{Y}] = \begin{cases} 1, & \text{if } 2Y_1 + \frac{1}{2}Y_2 > 0 \\ -1, & \text{otherwise.} \end{cases}$$

b) By symmetry,

$$P(\hat{X} \neq X) = P[\hat{X} = +1|X = -1] = P[2Y_1 + \frac{1}{2}Y_2 > 0|X = -1].$$

Now, if $X = -1$, $2Y_1 + \frac{1}{2}Y_2 = N(-2.5, 4 + \frac{1}{4} \times 4 = 5)$. Also,

$$P(N(-2.5, 5) > 0) = P(N(0, 1) > \frac{2.5}{\sqrt{5}}) = P(N(0, 1) > 1.12) \approx 0.13,$$

according to the values given in the table.

6. (15%)

Can you find two random variables X, Y such that $E[X|Y] > Y$ and $E[Y|X] > X$? Either give an example or prove that it is not possible.

We cannot find such random variables because $E[X|Y] > Y$ implies, by taking expectation, $E(X) > E(Y)$ whereas $E[Y|X] > X$ implies $E(Y) > E(X)$.